Am. J. Hum. Genet. 67:1428–1436, 2000

Diagnostic Testing for Rett Syndrome by DHPLC and Direct Sequencing Analysis of the *MECP2* Gene: Identification of Several Novel Mutations and Polymorphisms

Inge M. Buyse,[1] Ping Fang,[1] Katherine T. Hoon,[1] Ruthie E. Amir,[1] Huda Y. Zoghbi,[1,2,3] and Benjamin B. Roa[1]

Departments of [1]Molecular and Human Genetics and [2]Pediatrics and [3]Howard Hughes Medical Institute, Baylor College of Medicine, Houston

Rett syndrome (RTT) is an X-linked dominant neurodevelopmental disorder affecting 1/10,000–15,000 girls. The disease-causing gene was identified as *MECP2* on chromosome Xq28, and mutations have been found in ~80% of patients diagnosed with RTT. Numerous mutations have been identified in de novo and rare familial cases, and they occur primarily in the methyl-CpG–binding and transcriptional-repression domains of MeCP2. Our first diagnostic strategy used bidirectional sequencing of the entire *MECP2* coding region. Subsequently, we implemented a two-tiered strategy that used denaturing high-performance liquid chromatography (DHPLC) for initial screening of nucleotide variants, followed by confirmatory sequencing analysis. If a definite mutation was not identified, then the entire *MECP2* coding region was sequenced, to reduce the risk of false negatives. Collectively, we tested 228 unrelated female patients with a diagnosis of possible (209) or classic (19) RTT and found *MECP2* mutations in 83 (40%) of 209 and 16 (84%) of 19 of the patients, respectively. Thirty-two different mutations were identified (8 missense, 9 nonsense, 1 splice site, and 14 frameshifts), of which 12 are novel and 9 recurrent in unrelated patients. Seven unclassified variants and eight polymorphisms were detected in 228 probands. Interestingly, we found that T203M, previously reported as a missense mutation in an autistic patient, is actually a benign polymorphism, according to parental analysis performed in a second case identified in this study. These findings highlight the complexities of missense variant interpretation and emphasize the importance of parental DNA analysis for establishing an etiologic relation between a possible mutation and disease. Overall, we found a 98.8% concordance rate between DHPLC and sequence analyses. One mutation initially missed by the DHPLC screening was identified by sequencing. Modified conditions subsequently enabled its detection, underscoring the need for multiple optimized conditions for DHPLC analysis. We conclude that this two-tiered approach provides a sensitive, robust, and efficient strategy for RTT molecular diagnosis.

Introduction

Rett syndrome (RTT [MIM 312750]) is an X-linked dominant neurodevelopmental disorder and a significant genetic cause of mental retardation, affecting 1/10,000–15,000 girls (Rett 1966; Hagberg et al. 1983; Hagberg 1985; Schanen and Francke 1998). Patients with classic RTT appear to develop normally until age 6–18 mo, at which time they enter a period of neurodevelopmental regression. Symptoms include gradual loss of speech and purposeful hand use, development of microcephaly, seizures, ataxia, autistic features, intermittent hyperventilation, and stereotypic hand movements. The condition subsequently stabilizes, with patients usually surviving into adulthood (Hagberg et al. 1983). The RTT gene on Xq28 was identified as *MECP2*, which encodes the methyl-CpG–binding protein 2 that is normally involved in transcriptional silencing (Amir et al. 1999 [GenBank accession number X99686]). Numerous studies have since found various mutations (missense, nonsense, and frameshifts) in the coding region of *MECP2* in patients with RTT, identifying mutations for as many as 80% of patients (Amir et al. 1999, 2000; Wan et al. 1999; Amano et al. 2000; Bienvenu et al. 2000; Cheadle et al. 2000; Hampson et al. 2000; Huppke et al. 2000; Kim and Cook 2000; Obata et al. 2000; Orrico et al. 2000; Xiang et al. 2000); the remaining 20% may have mutations in other regions of this gene, such as regulatory elements and noncoding regions, but this remains to be determined. Most mutations lie within the methyl-CpG–binding (MBD) or transcriptional-repression (TRD) functional domains. In addition, a number of deletions

Received September 13, 2000; accepted for publication October 16, 2000; electronically published October 30, 2000.

Address for correspondence and reprints: Dr. Benjamin B. Roa, Baylor DNA Diagnostic Laboratory, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Room 205D, Houston, TX 77030. E-mail: broa@bcm.tmc.edu

0002-9297/2000/6706-0009$02.00

Table 1

Primer Sequences Used for PCR and Dye-Terminator Sequencing

Exon	Sequence	Fragment Size (bp)
PCR primers:		
2-For	U-TAA GCT GGG AAA TAG CCT AGT AC	454
2-Rev	R-TTA TAT GGC ACA GTT TGG CAC AG	
3-For	U-AGG ACA TCA AGA TCT GAG TGT AT	662
3-Rev	R-GGT CAT TTC AAG CAC ACC TG	
4a-For	U-CGA GTG AGT GGC TTT GGT GA	715
4a-Rev	R-ACA GAT CGG ATA GAA GAC TCC TT	
4a-For.2	U-CGC TCT GCC CTA TCT CTG A	620
4b-For.3	U-GGC AGG AAG CGA AAA GCT GAG	402
4b-Rev.3	R- TGA GTG GTG GTG ATG GTG GTG G	
4c/d-cFor	U-GGA AAG GAC TGA AGA CCT GTA AG	408
4c/d-dRev	R-CTC CCT CCC CTC GGT GTT TG	
4e-For	U-GGA GAA GAT GCC CAG AGG AG	447
4e-Rev	R-CGG TAA GAA AAA CAT CCC CAA	
U (−21 M13 primer tail)	TGT AAA ACG ACG GCC AGT	
R (M13 reverse tail)	CAG GAA ACA GCT ATG ACC	
Dye-terminator sequencing primers:		
2-Rev.2	CTA AAA AAA AAA AAA GGA AGG TTAC	
4c-For.S	AGC CCT GGG CGG AAA AGC	
4d-Rev.S	TAC TTT TCT GCG GCC GTG	

have been identified within a limited region, which suggests a potential recombinational hotspot in the last exon of the gene. A number of mutations recur in unrelated patients, reflecting the hypermutability of certain sites. It is noteworthy that *MECP2* mutations have also been found to cause a variety of phenotypes other than classic RTT, ranging from very mild mental retardation to severe neonatal encephalopathy in boys born into kindreds with RTT (Wan et al. 1999; Orrico et al. 2000).

Many methods for detecting mutations have been described, and strengths and limitations are inherent in each technique (Cotton 1997). Because DNA-sequence analysis is considered to be the gold standard for the identification of point mutations or deletion/insertion mutations that involve a few bases, DNA diagnostic testing was initially set up by PCR-based direct sequencing of the *MECP2* coding region, using automated fluorescence methods. Nevertheless, we thought that the RTT diagnostic test could be made more efficient by use of a robust method to scan patient samples for sequence variations/mutations prior to targeted sequence analysis. Denaturing high-performance liquid chromatography (DHPLC) is such a method. A highly sensitive PCR-based technique for nucleotide variant detection, DHPLC relies on the principle of heteroduplex analysis by ion-pair reverse-phase liquid chromatography under partially denaturing conditions (Oefner and Underhill 1995, 1998; Liu et al. 1998; O'Donovan et al. 1998). We thus devised a two-tiered molecular diagnostic approach for RTT, to increase test efficiency while maintaining the sensitivity provided by sequence analysis. The technical strategy and our cumulative results are presented in this article.

Material and Methods

Patient Material

Genomic DNA from patients with RTT with a previously identified mutation in the *MECP2* coding region was used as positive control material for the development of DNA diagnostic tests (Amir et al. 1999, 2000; R. E. Amir and H. Y. Zoghbi, unpublished data). We tested 209 female patients with possible RTT and 19 with a diagnosis of classic RTT (as determined by the referring clinicians), whose blood samples were submitted to the Baylor DNA Diagnostic Laboratory. Genomic DNA was extracted from blood leukocytes using the Puregene DNA isolation kit (Gentra Systems) or the QIAamp DNA Blood kit (Qiagen), following the manufacturer's instructions.

PCR Amplification

PCR primers (table 1) were designed to amplify *MECP2* coding exons 2, 3, and 4, using a total of six reactions (these were exons 1, 2, and 3 before the recent discovery of a new 5′ UTR exon [Reichwald et al. 2000]). Primers for coding exons 2 and 3 correspond to flanking intron sequences. Exon 4 was amplified as four overlapping fragments (4a, 4b, 4cd, and 4e) that collectively span the 5′ intronic sequence and 3′ UTR sequences. Two forward primers were used to amplify exon 4a: the exon

4a-For primer (used for sequencing) was redesigned as exon 4a-For.2 (used for DHPLC), to prevent upstream polymorphisms from interfering with DHPLC analysis. Both primers are used in combination with the exon 4a-Rev primer. Primers (Gibco BRL) were synthesized with universal M13 tails (−21M13 or M13 reverse), to facilitate direct sequencing using dye-primer chemistry (see table 1). PCR reactions were carried out in 50-μl reaction volumes containing 100 ng genomic DNA, 1 × PCR buffer (50 mM KCl, 10 mM Tris HCl [pH 8.3], 1.5 mM $MgCl_2$, and 0.001% w/v gelatin; Perkin Elmer), 0.05 mM dNTP, 1.88 pmol of each primer, and 1.25 U *Taq* polymerase (Perkin Elmer). The exon 4c/d PCR reaction mix contained 1 mM $MgCl_2$ and 4.69 pmol of each primer. PCR conditions included an initial denaturation at 94°C for 2 min 30 s, followed by 10 "step-down" cycles of 30 s at 94°C, 30 s at 65°C (decreasing 1.5°C per cycle), and 1 min 45 s at 72°C, followed by 28 cycles of 30 s at 94°C, 30 s at 51°C, and 1 min 30 s at 72°C, and a final extension step at 72°C for 5 min.

DHPLC Analysis

Heteroduplex formation was induced by heat denaturation of PCR products at 94°C for 5 min, followed by gradual reannealing from 94°C to 25°C over 45 min. DHPLC analysis was performed with the WAVE DNA-fragment analysis system (Transgenomic). PCR products (10 μl per sample) were eluted at a flow rate of 0.9 ml/min with a linear acetonitrile gradient. The values of the buffer gradients (buffer A, 0.1 M triethylammoniumacetate; buffer B, 0.1M triethylammoniumacetate/25% acetonitrile), start and end points of the gradient, and melting temperature predictions were determined by WAVEMAKER software (Transgenomic). Analysis per sample took ~7.5 min, including regeneration and reequilibration to the starting conditions. Optimal run temperatures were empirically determined; mobile-phase temperatures were assessed within a 5°C window above and below the suggested run temperature on the basis of each fragment's characteristic melting profile. Run temperatures that allowed detection of all tested sequence variants were 59°C for exon 2; 61°C, 63°C, 66°C, and 67°C for exon 3; 61°C, 64°C, and 66°C for exon 4a; 64°C and 65°C for exon 4b; 65°C and 66°C for exon 4c/d; and 60°C, 63°C, and 65°C for exon 4e. Data analysis was based on visual inspection of the chromatograms, and comparison with normal controls was included in each run. Heterozygous profiles were detected as distinct elution peaks from homozygous wild-type peaks.

Direct Sequencing Analysis

PCR products used for sequencing analysis were purified using the QIAquick PCR purification kit (Qiagen) and bidirectionally sequenced using the ABI Prism BigDye Primer Cycle Sequencing Ready Reaction kit (PE Biosystems). The BigDye Terminator Cycle Sequencing Ready Reaction kit (PE Biosystems) was used to sequence the exon 2 reverse reaction and the exon 4c/d forward and reverse reactions (primers listed in table 1). Samples were analyzed on an ABI 377 DNA sequencer (PE Biosystems), according to the manufacturer's instructions. Patient sequence data from both orientations were aligned for comparison with corresponding wild-type sequence using SEQUENCHER 3.0 analysis software.

Our two-step diagnostic strategy calls for targeted sequence analysis of a specific region for samples that are positive by DHPLC analysis. For samples that are negative by DHPLC analysis or are not found to have a definitive mutation by targeted sequence analysis (polymorphism or unclassified variant alleles), direct sequencing of the entire *MECP2* coding region is performed, to reduce the risk of false negatives.

Results

MECP2 *Mutation Detection by Direct Sequence Analysis*

Mutation analysis for RTT was initially set up using bidirectional sequencing of PCR products corresponding to the *MECP2* coding region. Dye-primer sequencing chemistry was used, except for several dye-terminator sequencing reactions required for technical reasons (exon 2 reverse and exon 4c/d forward and reverse reactions). Control samples used in our assay validation included 11 previously characterized DNA samples from patients with a diagnosis of classic RTT and from unaffected family members (Amir et al. 1999, 2000; R. E. Amir and H. Y. Zoghbi, unpublished data). Sequence analysis, in accordance with our protocol, was performed in a blinded manner, and 11 of 11 control samples were correctly identified as mutant, polymorphic, or negative.

Diagnostic sequencing was performed on the first 142 patients referred to the Baylor DNA Diagnostic Laboratory with a definite or possible diagnosis of RTT. Sequence variations were observed in 65 of 142 patients. Of these, 62 (44%) were positive for a disease-causing *MECP2* mutation. A mutation was considered disease-causing under either of the following conditions: (1) it had already been reported in the literature as a de novo mutation in a patient with RTT or (2) it was a truncating mutation that disrupted gene function (nonsense, splice-site mutation, insertion, or deletion frameshift). In three individuals (2%) who were heterozygous for an unclassified sequence variant, analysis of both parents was recommended to define each variant as either a de novo

mutation or a benign polymorphism. The remaining 77 cases (54%) were negative by sequencing (see table 2).

MECP2 *Mutation Detection by DHPLC Analysis*

We evaluated DHPLC for its potential as a screening method to reduce the need for sequencing the complete coding region in almost half of our caseload. PCR heteroduplexes are resolved from homoduplexes on a DHPLC column via differential elution profiles under partially denaturing conditions. We optimized DHPLC run conditions with the aid of WAVEMAKER software and by empiric determination using 50 positive control samples (see Material and Methods section) that included mutations (base substitutions, insertions, or deletions), polymorphisms, and unclassified missense variants in exons 3 and 4. Because no exon 2 mutations have been identified at this institution to date, exon 2 run conditions were based on software prediction. Exons 3 and 4 contain multiple melting domains, so multiple run temperatures were used to analyze PCR fragments in these regions. All 50 sequence variants were identified, under one or more run conditions, as unique elution profiles. Examples of variant DHPLC chromatograms and their confirmatory sequences are shown in figure 1.

The validation of *MECP2* coding-region analysis by DHPLC consisted of two phases. For the first phase, a set of 15 samples that were previously tested by sequence analysis were analyzed by DHPLC in a blinded manner. DHPLC analysis of the entire *MECP2* coding yielded 100% concordance with prior sequencing data (10 positives and 5 negatives; see table 2). In the second phase of DHPLC validation, 36 samples that were being examined by sequence analysis in our laboratory were tested in parallel by DHPLC. Nineteen samples were found to carry one or more sequence variations and 17 were negative, which yielded 100% concordance between both methods (table 2).

On the basis of these results, a two-tiered molecular diagnostic strategy was adopted. Our current testing strategy entails that we first analyze all *MECP2* coding exons by DHPLC. PCR fragments encoding a sequence variant are further analyzed by bidirectional sequencing. For samples that are negative by initial DHPLC analysis or are found to carry a polymorphism or unclassified sequence variant, the entire *MECP2* coding region is analyzed by bidirectional sequencing. This strategy proved to be both efficient and robust. Eighty-six cases

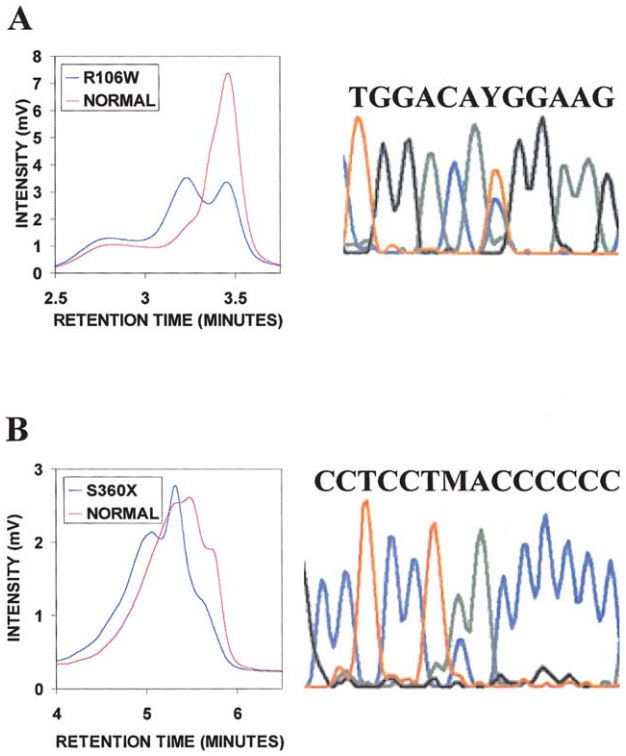

Figure 1 DHPLC elution profiles for two *MECP2* coding-region mutations detected in two patients with sporadic RTT. *A*, Relevant exon 3 elution profiles (63°C) for a normal individual and a patient carrying the R106W mutation. The DHPLC profiles are superimposed to illustrate differences. The direct sequencing result illustrates the corresponding 316C→T nucleotide substitution. *B*, Relevant exon 4b elution profiles (64°C) for a normal individual and a patient carrying the S360X mutation. The direct sequencing result illustrates the corresponding 1079C→A nucleotide substitution.

Table 2

Different Phases Involved in the Development of a Two-Step Protocol for RTT Testing by DHPLC and Bidirectional Direct Sequencing Analysis

	No. (%)		
Test Data	Positive	Unclassified	Negative
Bidirectional direct sequencing (142 cases)	62 (44)	3 (2)	77 (54)
DHPLC analysis validation:			
Validation phase 1 (15 samples tested blindly[a])	10		5
Validation phase 2 (36 cases tested in parallel[a])	19		17
DHPLC and bidirectional direct sequencing (86 cases[b])	37 (43)	2 (2)	47 (55)

[a] 100% concordance.
[b] 98.8% concordance: one case DHPLC negative and sequencing positive.

Table 3

***MECP2* Mutations Detected by DHPLC and Direct Sequencing Analysis**

Mutation Type and Exon	Nucleotide Change	Amino Acid Change	Domain	No. of Patients (This Lab)	Reference
Missense:					
3	317 G→A	R106Q	MBD	2	Bienvenu et al. (2000)
3	316 C→T	R106W	MBD	3	Amir et al. (1999)
4	397 C→T	R133C	MBD	1	Amir et al. (1999)
4	455 C→G	P152R	MBD	1	Cheadle et al. (2000)
4	464 T→C	F155S	MBD	1	Amir et al. (1999)
4	473 C→T	T158M	MBD	21	Amir et al. (1999)
4	916 C→T	R306C	TRD	8	Wan et al. (1999)
4	917 G→A	R306H	TRD	1	Cheadle et al. (2000)
Nonsense:					
4	423 C→G	Y141X	MBD	1	Amir et al. (2000)
4	430 A→T	K144X	MBD	1	This study
4	502 C→T	R168X		13	Wan et al. (1999)
4	508 C→T	Q170X		1	Amir et al. (2000)
4	611 C→G	S204X		1	This study
4	763 C→T	R255X	TRD	12	Amir et al. (1999)
4	808 C→T	R270X	TRD	8	Cheadle et al. (2000)
4	880 C→T	R294X	TRD	7	Cheadle et al. (2000); Xiang et al. (2000)
4	1079 C→A	S360X		1	This study
Splice site	IVS2-2 A→G			1	Huppke et al. (2000); Amir et al. (2000)
Frameshift:					
4	553delG			1	This study
4	705delG		TRD	1	Amir et al. (2000)
4	750delC		TRD	1	This study
4	749insCC		TRD	1	This study
4	803delG		TRD	2	Wan et al. (1999)
4	807delC		TRD	1	Obata et al. (2000)
4	965del6nt/ 1026insG/ 1138del71nt	Complex deletion/ insertion		1	This study
4	1118del183nt/ ins61nt	Complex deletion/ insertion		1	This study
4	1157del41nt			1	Cheadle et al. (2000)
4	1161del6nt/ 1177del26nt	Complex deletion		1	This study
4	1163del26nt			1	Bienvenu et al. (2000)
4	1163del11nt/ 1174del18nt	Complex deletion		1	This study
4	1164del44nt			1	This study
4	1308delTC			1	This study

have been analyzed using this strategy (see table 2). Mutations were identified in 37 cases (43%), 2 cases (2%) had unclassified variants, and 47 cases (55%) were negative. The DHPLC results were consistent with sequencing analysis in 98.8% of these cases. One patient was initially negative by DHPLC analysis, but direct sequencing of the complete *MECP2* coding region of this patient revealed an unclassified missense substitution in exon 3 (S86C). This substitution was missed by DHPLC, despite the use of three different temperatures (61°C, 63°C, and 66°C), causing a false negative rate of 1.2%. The region encoding amino acids 85–90 is very GC rich but was anticipated to melt at 66°C. Reanalysis by DHPLC allowed detection of this specific variant at 67°C (but not at 66°C). This temperature was subsequently added to the current set of running conditions (see the Material and Methods section).

MECP2 *Sequence Variations Detected*

Our laboratory tested a total of 228 unrelated female patients with a diagnosis of possible (209) or classic (19) RTT for *MECP2* mutations. Table 3 lists all the identified mutations, and table 4 lists the polymorphic and unclassified sequence variations detected in this group of patients. Disease-causing mutations were detected in 83 (40%) of 209 and 16 (84%) of 19 patients with possible and classic sporadic RTT, respectively. A total of 32 different mutations are reported, of which 12 are novel (3 nonsense and 9 frameshift mutations; table 3). These novel frameshifts included two complex deletion and two complex deletion/insertion mutations. These mutations were verified by comparison of forward and reverse strands, and sequence analysis of the mutant allele exclusively in cases of large deletions. Nine muta-

tions recurred in 77% of our patients bearing a *MECP2* mutation (table 3).

A total of eight polymorphisms (five silent and three missense) were detected in 15 patients (table 4). S194, S411, and A444T appear to be more common, recurring two, six, and two times, respectively. Of 15 patients with polymorphisms, 12 were found to carry a disease-causing mutation together with the polymorphism, although the chromosomal phase was not set. Seven unclassified sequence variants were found, for which parental analysis was recommended. Although parental samples were not available, two of these unclassified variants are consistent with being polymorphisms (S291A and V412I), since they were identified in patients who also had encoded a classified *MECP2* mutation.

Parental analysis performed in two cases strongly indicates that the T203M and A444T alleles are benign polymorphisms, because the unaffected father carried the missense variant found in the affected child. Of particular interest is one case in which the patient was found to carry both the R106Q and T203M variant alleles. Because neither missense allele had been reported in the literature at the time, we requested parental samples for further analysis, to help classify the alleles as mutant or polymorphic. The R106Q allele was not found in either unaffected parent and was interpreted to be a de novo mutation. Our findings were subsequently confirmed by a published report of another de novo R106Q mutation in a typical case of RTT (Bienvenu et al. 2000). In contrast, our parental analysis of the T203M allele was consistent with a benign polymorphism. Figure 2 shows the sequencing data for this family, which demonstrate that T203M is present in both the proband and her clinically unaffected father.

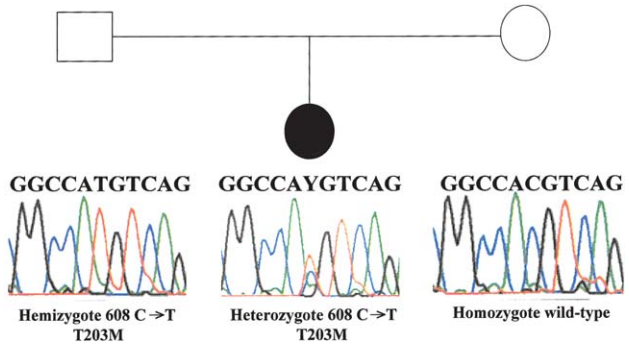

Figure 2 Familial sequence analysis for the T203M polymorphism. Direct-sequencing results illustrate the presence of the 608C→T nucleotide substitution in the affected patient (heterozygous) and in her clinically unaffected father (hemizygous) but not in her unaffected mother.

Prenatal Diagnosis

Our laboratory has performed four prenatal tests thus far. A mutation (R106W, P152R, R168X, and R294X) was identified in the index case for each family. Subsequent analysis of maternal DNA by DHPLC and by direct sequencing of the PCR fragment of interest suggested that the mutation arose de novo in each case, although germline mosaicism was not excluded. Prenatal diagnosis by DHPLC and direct sequencing performed on amniotic fluid and cultured amniocytes was negative

Table 4

***MECP2* Polymorphisms and Unclassified Sequence Variants Detected by DHPLC and Direct Sequencing Analysis**

Variant and Exon	Nucleotide Change	Amino Acid Change	Domain	No. of Patients (This Lab)	Reference
Polymorphism:					
3	375 C→A	I125	MBD	1	Cheadle et al. (2000)
4	582 C→T	S194		2	Amir et al. (2000)
4	608 C→T	T203M		1	Orrico et al. (2000)
4	843 C→T	A281	TRD	1	This study
4	984 C→T	L328		1	This study
4	1189 G→A	E397K		1	Wan et al. (1999)
4	1233 C→T	S411		6	Amir et al. (1999)
4	1330 G→A	A444T		2	This study
Unclassified:					
3	257 C→G	S86C	MBD	1	This study
3	298 C→G	L100V	MBD	1	This study
4	859 G→C	A287P	TRD	1	This study
4	871 T→G	S291A[a]	TRD	1	This study
4	914 A→G	K305R	TRD	1	This study
4	1234 G→A	V412I[a]		1	This study
4	1164del9nt (in-frame del)			1	This study

[a] Probable polymorphism, occurring in a patient who also carried a classified *MECP2* mutation.

for the mutation in all four cases. Maternal cell contamination was ruled out by PCR analysis of short-tandem repeats at other loci.

Discussion

This article presents the mutation data accumulated from our diagnostic laboratory experience, which includes testing of 228 unrelated patients with a diagnosis of possible (209) or classic (19) RTT. Disease-causing mutations were detected in 84% of patients with classical sporadic RTT, which is consistent with the estimate reported in the literature. That only 40% of the patients suggestive for RTT were positive for *MECP2* mutations reflects the clinical heterogeneity of these patients, who were referred from different sources. Since mutation analysis has focused largely on the *MECP2* coding region, any large rearrangements and/or mutations that might exist in the regulatory and noncoding regions are not addressed. Nonetheless, the existing data from research and diagnostic laboratories demonstrate that *MECP2* coding-region mutations account for the majority of classic RTT. We found a total of 32 different mutations (table 3), including 12 novel *MECP2* coding-region mutations. Nine of these are novel frameshift mutations, with the majority located in the last exon. Furthermore, four of these novel frameshifts involve complex insertion and/or deletion mutations in exon 4. These findings are consistent with the region being a recombinational hotspot containing palindromic and quasipalindromic sequences (Amano et al. 2000; Amir et al. 2000; Bienvenu et al. 2000; Cheadle et al. 2000; Huppke et al. 2000; Xiang et al. 2000). We identified nine mutations that recurred in 77% of our mutation-positive patients (table 3). Seven of these recurrent mutations (R106W, T158M, R306C, R168X, R255X, R270X, and R294X) involve C→T transitions at CpG dinucleotides (Wan et al. 1999; Amano et al. 2000; Amir et al. 2000; Bienvenu et al. 2000; Cheadle et al. 2000; Hampson et al. 2000; Huppke et al. 2000; Obata et al. 2000). We also found eight *MECP2* sequence polymorphisms and seven unclassified *MECP2* missense variants for which parental analyses were recommended (table 4).

Parental analysis strongly suggested that two missense variants identified in this study, A444T and T203M, are polymorphisms; the variant allele was found in the clinically unaffected father. Our findings are particularly relevant for the T203M allele, which was recently reported as a disease-causing mutation in a girl who was initially diagnosed with autism and subsequently characterized as having variant RTT with autistic features (Orrico et al. 2000). This patient had been adopted, so parental analysis could not be performed; instead, analysis was performed on 200 normal chromosomes, which were apparently negative for T203M. The authors concluded that T203M is a novel missense mutation associated with an atypical RTT phenotype (Orrico et al. 2000). In retrospect, the fact that the threonine 203 residue is not evolutionarily conserved in MeCP2 provided a clue that this could be a benign polymorphism. Moreover, our patient with the T203M polymorphism was also found to carry a de novo R106Q mutation (table 3; Bienvenu et al. 2000). Our collective findings reclassify T203M as a polymorphism and highlight the need for caution in the clinical interpretation of sequence variants, such as missense alleles, that may or may not be disease associated. T203M appears to be a rare polymorphic allele, since we did not find it in the other 455 patient chromosomes we tested, and it was not found in 200 normal chromosomes (Orrico et al. 2000). Further molecular investigation, including parental analysis, would be required for allele classification in such cases.

This two-tiered approach has a number of advantages over our original sequencing protocol. It is less labor and reagent intensive than fluorescent gel sequencing, and testing efficiency is increased by prescreening patient samples by DHPLC prior to targeted sequence analysis (the amount of sequencing was reduced by a factor of six for the 40% of cases in which we detected mutations). On the basis of our experience, the operational cost of *MECP2* mutation analysis by this combined strategy is ~14% lower than full-sequence analysis. Comparisons were based on reagent and labor costs specific to each assay and did not include the costs of equipment, repeat testing, or sample processing common to both assays (e.g., DNA preparation and laboratory overhead). On the basis of our detection rate and sample volumes over a 1-year period, reagent costs were estimated to be 9% less for the combined DHPLC and targeted-sequencing strategy than those for full-sequence analysis alone. Personnel efficiency is also improved: our combined approach requires 26 technician h/wk versus 32 h for full sequencing of an average weekly caseload. This translates into a 19% reduction in labor, or almost 1 technician d/wk. The actual cost per test will vary given the different sample volumes, rate of repeat testing, and overhead cost structures among laboratories. Apart from increased cost efficiency, the diagnostic testing strategy combining DHPLC and direct sequencing has proved to be a sensitive and efficient method for *MECP2* mutation analysis. The sensitivity of this combined approach is equal to or greater than that of sequencing. Mutation-positive samples were initially identified by DHPLC in all but one case (see table 2). Any variants that could be missed by DHPLC would be identified by sequence analysis of

the complete *MECP2* coding region, which is prescribed for all samples that test negative by DHPLC. Likewise, samples that test positive by DHPLC and are found to have a polymorphism or unclassified variant by targeted sequencing are subsequently sequenced for the complete coding region, further reducing the risk for false negatives. Negative samples benefit from routine analysis by two sensitive and independent methods, so the sensitivity of the combined approach may be greater than that of sequencing alone. The collective data from our first 86 cases demonstrate the strength of this approach.

Further advantages of DHPLC include computer-assisted determination of analytical conditions and automated sample-handling features. Nevertheless, mutation detection is dependent on the complexity of each fragment's sequence-specific melting profile and the optimization of DHPLC analytical conditions for each specific fragment. This fact was brought to the forefront by our experience with one patient, who tested negative by DHPLC analysis but was positive for a novel missense variant (S86C) by sequencing, despite the use of three different DHPLC temperature conditions for that fragment. Temperature conditions had been determined by a combination of computer software predictions and empirical data that used available positive controls. Although four different variants were tested during development of the DHPLC run conditions for exon 3 (L100V, R106W, R106Q, and I125), these were located slightly downstream of S86C. Nevertheless, software predictions suggested that sequence alterations in this region would be detected. When repeat analysis of this patient's sample revealed that 67°C rather than 66°C allowed detection of the S86C variant, we added this temperature to the DHPLC run conditions.

Improvements in the melting-profile software should increase sensitivity and facilitate the use of DHPLC for diagnostic testing of unknown mutations in any given region of interest. DHPLC analysis may not detect homozygous or hemizygous point mutations without mixing equal amounts of test and control samples to induce heteroduplex formation. This would mean that samples from boys with possible atypical RTT should be sequenced entirely, a policy currently followed in this laboratory. In sum, we conclude that the use of DHPLC as an initial screening tool is ideal for *MECP2* mutation analysis. The two-tiered strategy combining DHPLC with direct sequence analysis provides a robust and efficient means of molecular diagnostic testing for RTT. No current analytical method, including sequence analysis, has 100% detection under all conditions. The challenge for molecular diagnostics is to maximize assay sensitivity, which can be accomplished by optimizing and combining appropriate analytical methods.

Acknowledgments

The authors wish to thank Dr. Igna Van den Veyver, Dr. Dani Bercovich, and Dr. Karla Bowles for thoughtful discussions and a critical review of this manuscript.

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

GenBank, http://www.ncbi.nlm.nih.gov/Genbank/index.html (for the *MECP2* gene [accession number X99686])
Online Mendelian Inheritance of Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for RTT [MIM 312750])

References

Amano K, Nomura Y, Segawa M, Yamakawa K (2000) Mutational analysis of the *MECP2* gene in Japanese patients with Rett syndrome. J Hum Genet 45:231–236

Amir RE, Van den Veyver IB, Schultz R, Malicki DM, Tran CQ, Dahle EJ, Philippi A, Timar L, Percy AK, Motil KJ, Lichtarge O, O'Brian Smith E, Glaze DG, Zoghbi HY (2000) Influence of mutation type and X chromosome inactivation on Rett syndrome phenotypes. Ann Neurol 47:670–679

Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY (1999) Rett syndrome is caused by mutations in X-linked *MECP2,* encoding methyl-CpG-binding protein 2. Nat Genet 23:185–188

Bienvenu T, Carrié A, de Roux N, Vinet M-C, Jonveaux P, Couvert P, Villard L, Arzimanoglou A, Beldjord C, Fontes M, Tardieu M, Chelly J (2000) *MECP2* mutations account for most cases of typical forms of Rett syndrome. Hum Mol Genet 9:1377–1384

Cheadle JP, Gill H, Fleming N, Maynard J, Kerr A, Leonard H, Krawczak M, Cooper DN, Lynch S, Thomas N, Hughes H, Hulten M, Ravine D, Sampson JR, Clarke A (2000) Long-read sequence analysis of the *MECP2* gene in Rett syndrome patients: correlation of disease severity with mutation type and location. Hum Mol Genet 9:1119–1129

Cotton RGH (1997) Slowly but surely towards better scanning for mutations. Trends Genet 13:43–46

Hagberg G (1985) Rett's syndrome: prevalence and impact on progressive severe mental retardation in girls. Acta Paediatr Scand 74:405–408

Hagberg B, Aicardi J, Dias K, Ramos O (1983) A progressive syndrome of autism, dementia, ataxia and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. Ann Neurol 14:471–479

Hampson K, Woods CG, Latip F, Webb T (2000) Mutations in the *MECP2* gene in a cohort of girls with Rett syndrome. J Med Genet 37:610–612

Huppke P, Laccone F, Krämer N, Engel W, Hanefeld F (2000) Rett syndrome: analysis of *MECP2* and clinical characterization of 31 patients. Hum Mol Genet 9:1369–1375

Kim S-J, Cook EH (2000) Novel *de novo* nonsense mutation

of *MECP2* in a patient with Rett syndrome. Hum Mutat 15:382–383

Liu W, Smith DI, Rechtzigel KJ, Thibodeau SN, James CD (1998) Denaturing high-performance liquid chromatography (DHPLC) used in the detection of germline and somatic mutations. Nucleic Acids Res 26:1396–1400

Obata K, Matsuishi T, Yamashita Y, Fukuda T, Kuwajima K, Horiuchi I, Nagamitsu S, Iwanaga R, Kimura A, Omori I, Endo S, Mori K, Kondo I (2000) Mutation analysis of the methyl-CpG-binding protein 2 gene (*MECP2*) in patients with Rett syndrome. J Med Genet 37:608–610

O'Donovan MC, Oefner PJ, Roberts SC, Ausin J, Hoogendoorn B, Guy C, Speight G, Upadhyaya M, Sommer SS, McMuffin P (1998) Blind analysis of denaturing high-performance liquid chromatography as a tool for mutation detection. Genomics 52:44–49

Oefner PJ, Underhill PA (1995) Comparative DNA sequencing by denaturing high-performance liquid chromatography (DHPLC). Am J Hum Genet 57:A266

——— (1998) DNA mutation detection using denaturing high-performance liquid chromatography (DHPLC). In: Dracopoli NC, Haines J, Korf BR, Morton C, Seidman CE, Seidman JG, Moir DT, Smith DR (eds) Current protocols in human genetics, Suppl 19. Wiley & Sons, New York, pp 7.10.1–7.10.12

Orrico A, Lam C-W, Galli L, Dotti MT, Hayek G, Tong S-F, Poon PMK, Zappella M, Federico A, Sorrentino V (2000) *MECP2* mutation in male patients with non-specific X-linked mental retardation. FEBS Lett 24106:1–4

Reichwald K, Thiesen J, Wiehe T, Weitzel J, Strätling WH, Kioschis P, Poustka A, Rosenthal A, Platzer M (2000) Comparative sequence analysis of the *MECP2*-locus in human and mouse reveals new transcribed regions. Mamm Genome 11:182–190

Rett A (1966) Über ein eigenartiges hirnatrophisches syndrom bei hyperamonaemie im kindesalter. Wien Med Wochenschr 116:723–728

Schanen NC and Francke U (1998) A severely affected male born into a Rett syndrome kindred supports X-linked inheritance and allows extension of the exclusion map. Am J Hum Genet 63:267–269

Wan M, Lee SSJ, Zhang X, Houwink-Manville I, Song H-R, Amir RE, Budden S, Naidu SB, Pereira JLL, Lo IFM, Zoghbi HY, Schanen NC, Francke U (1999) Rett syndrome and beyond: recurrent spontaneous and familial *MECP2* mutations at CpG hotspots. Am J Hum Genet 65:1520–1529

Xiang F, Buervenich S, Nicolao P, Bailey MES, Zhang Z, Anvret M (2000) Mutation screening in Rett syndrome patients. J Med Genet 37:250–255